© 1998 Oxford University Press Human Molecular Genetics, 1998, Vol. 7, No. 3 507–515

Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan–Zonana syndrome, two hamartoma syndromes with germline *PTEN* mutation

Debbie J. Marsh[1], Valérie Coulon[2], Kathryn L. Lunetta[3], Philippe Rocca-Serra[2], Patricia L. M. Dahia[1], Zimu Zheng[1], Danny Liaw[4], Stacey Caron[1], Bernadette Duboué[2], Albert Y. Lin[5], Anne-Louise Richardson[6], Jean-Marie Bonnetblanc[7], Jean-Marie Bressieux[8], Agnés Cabarrot-Moreau[9], Agnés Chompret[10], Liliane Demange[11], Rosalind A. Eeles[12], Alan M. Yahanda[13], Eric R. Fearon[14], Jean-Pierre Fricker[15], Robert J. Gorlin[16], Shirley V. Hodgson[17], Susan Huson[18], Didier Lacombe[19], Frédéric LePrat[20], Sylvie Odent[21], Claude Toulouse[22], Olufunmilayo I. Olopade[23], Hagay Sobol[24], Sigrid Tishler[25], C. Geoffrey Woods[26], Bruce G. Robinson[6], H. Christian Weber[27], Ramon Parsons[4], Monica Peacocke[28], Michel Longy[2,*] and Charis Eng[1,29,*]

[1]Department of Adult Oncology and Charles A. Dana Human Cancer Genetics Unit, Dana-Farber Cancer Institute, Department of Medicine, Harvard Medical School, 1 Jimmy Fund Way, Boston, MA 02115-6084, USA, [2]Molecular Oncology Laboratory, Institut Bergoni, 180 rue de Saint-Gens, 33076 Bordeaux, France, [3]Department of Biostatistical Science, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA, [4]Departments of Pathology and Medicine, Columbia University Cancer Center, New York, NY 10032, USA, [5]Genetic Epidemiology Branch, National Cancer Institute, Bethesda, MD 20892, USA, [6]Kolling Institute of Medical Research, Royal North Shore Hospital and Department of Medicine, Sydney University, Sydney, NSW, Australia, [7]Department of Dermatology, CHRU Dupuytren, Limoges, France, [8]Department of Dermatology, CHU, Troyes, France, [9]Department of Dermatology, CHU Rangueil, Toulouse, France, [10]Department of Dermatology, Institut Gustave Roussy, Villejuif, France, [11]Oncogenetic Unit, Institut Jean Godinot, Reims, France, [12]Institute of Cancer Research, Royal Marsden Hospital, Sutton, Surrey SM2 5PT, UK, [13]Department of Surgery and [14]Division of Medical and Molecular Genetics, University of Michigan School of Medicine, Ann Arbor, MI, USA, [15]Oncogenetic Unit, Centre Paul Strauss, Strasbourg, France, [16]Department of Oral Pathology and Genetics, University of Minnesota, Minneapolis, MN, USA,[17]Clinical Genetics Department, Guys and St Thomas' Hospital, London SE1 9RT, UK, [18]Oxford Regional Genetics Service, Department of Clinical Genetics, Oxford Radcliffe Hospital Trust, The Churchill Hospital, Oxford OX3 7LJ, UK, [19]Department of Medical Genetics, CHU, Bordeaux, France, [20]Department of Endocrinology, CHU de Bordeaux, Pessac, France, [21]Department of Medical Genetic, CHRU de Rennes, Pontchaillou, France, [22]Oncogenetic Unit, Institut Bergoni, Bordeaux, France, [23]Section of Haematology/Oncology, Department of Medicine, University of Chicago Medical Centre, Chicago, IL 60637, USA, [24]Department of Oncogenetics, INSERM CRI 9703, Institut Paoli Calmettes, Marseille, France, [25]Haematology/Oncology, Harvard Pilgrim Health Plan, Boston, MA 02215, USA, [26]Department of Clinical Genetics, St James' University Hospital, Leeds LS9 7TF, UK, [27]Section of Gastroenterology, Boston University School of Medicine, Boston, MA 02118, USA, [28]Departments of Medicine and Dermatology, College of Physicians and Surgeons, Columbia University, New York, NY, USA and [29]CRC Human Cancer Genetics Research Group, University of Cambridge, Cambridge CB2 2QQ, UK

Received 18 October, 1997; Revised and Accepted 28 November, 1997

The tumour suppressor gene *PTEN*, which maps to 10q23.3 and encodes a 403 amino acid dual specificity phosphatase (protein tyrosine phosphatase; PTPase), was shown recently to play a broad role in human malignancy. Somatic *PTEN* deletions and mutations were observed in sporadic breast, brain, prostate and kidney cancer cell lines and in several primary tumours such as endometrial carcinomas, malignant melanoma

*To whom correspondence should be addressed at: Dana-Farber Cancer Institute, Charles A. Dana Human Cancer Genetics Unit, Richard and Susan Smith Laboratories, SM822, 1 Jimmy Fund Way, Boston, MA 02115, USA. Tel: +1 617 632 5632; Fax: +1 617 632 4280; Email: charis_eng@macmailgw.dfci.harvard.edu (C.E.) or Tel: +33 556 33 33 33; Fax: +33 556 33 33 89 (M.L.)

and thyroid tumours. In addition, *PTEN* was identified as the susceptibility gene for two hamartoma syndromes: Cowden disease (CD; MIM 158350) and Bannayan–Zonana (BZS) or Ruvalcaba–Riley–Smith syndrome (MIM 153480). Constitutive DNA from 37 CD families and seven BZS families was screened for germline *PTEN* mutations. *PTEN* mutations were identified in 30 of 37 (81%) CD families, including missense and nonsense point mutations, deletions, insertions, a deletion/insertion and splice site mutations. These mutations were scattered over the entire length of *PTEN*, with the exception of the first, fourth and last exons. A 'hot spot' for *PTEN* mutation in CD was identified in exon 5 that contains the PTPase core motif, with 13 of 30 (43%) CD mutations identified in this exon. Seven of 30 (23%) were within the core motif, the majority (five of seven) of which were missense mutations, possibly pointing to the functional significance of this region. Germline *PTEN* mutations were identified in four of seven (57%) BZS families studied. Interestingly, none of these mutations was observed in the PTPase core motif. It is also worthy of note that a single nonsense point mutation, R233X, was observed in the germline DNA from two unrelated CD families and one BZS family. Genotype–phenotype studies were not performed on this small group of BZS families. However, genotype–phenotype analysis in the group of CD families revealed two possible associations worthy of follow-up in independent analyses. The first was an association noted in the group of CD families with breast disease. A correlation was observed between the presence/absence of a *PTEN* mutation and the type of breast involvement (unaffected versus benign versus malignant). Specifically and more directly, an association was also observed between the presence of a *PTEN* mutation and malignant breast disease. Secondly, there appeared to be an inter-dependent association between mutations upstream and within the PTPase core motif, the core motif containing the majority of missense mutations, and the involvement of all major organ systems (central nervous system, thyroid, breast, skin and gastro-intestinal tract). However, these observations would need to be confirmed by studying a larger number of CD families.

INTRODUCTION

Hamartomas are developmentally disorganized, benign growths and are characteristic of two clinically related autosomal dominant syndromes, Cowden disease (CD) and Bannayan–Zonana (BZS) or Ruvalcaba–Riley–Smith syndrome. The predominant phenotype of CD, or multiple hamartoma syndrome, is hamartomas in many organ systems including the breast, thyroid, skin, central nervous system (CNS) and gastrointestinal (GI) tract (1–5). Trichilemmomas, benign tumours of the hair follicle infundibulum, and mucocutaneous papillomatosis are hallmarks of CD, occurring in 99% of CD patients (2–4). Other common hamartomas in CD patients include thyroid adenomas and multinodular goitres (40–60% of CD patients), breast fibroadenomas (70% of affected females) and hamartomatous GI polyps (35–40%). Affected CD individuals are at an increased risk of both breast and thyroid cancers, breast cancers developing in 25–50% of women and thyroid cancer in 3–10% of all affected individuals (4,5). CNS manifestations are present in this syndrome, occurring in ~40% of cases (4). Megencephaly/macrocephaly is present in 38% of CD patients. Lhermitte–Duclos disease (LDD), dysplastic gangliocytoma of the cerebellum, can lead to seizures, tremors and poor coordination, and is now established as a component of CD (6–9). CNS tumours, both benign and malignant, may also be components of CD.

Unlike CD, an increased risk of malignancy is not documented as being present in BZS. However, the partial clinical overlap between the two syndromes is exemplified by the presence of intestinal hamartomatous polyps, macrocephaly and multiple lipomas (occurring in a minority of CD patients). Other features that may be observed in BZS include an increased risk of intracranial tumours, vascular malformations, Hashimoto's thyroiditis, pigmented macules of the glans penis, delayed motor development, mild mental retardation and an early age of onset (10–13). Elucidation of a possible allelic relationship between these hamartoma syndromes awaited the isolation of a candidate susceptibility gene.

The CD susceptibility gene was localized to a 5 cM interval on chromosome sub-band 10q22–23 between the microsatellite markers *D10S541* and *D10S564* (14). Subsequently *PTEN* [alternatively named *MMAC1* and *TEP1* by independent groups (15,16)], a nine exon gene in this region encoding a dual specificity phosphatase with homology to tensin and auxilin, was identified as the CD susceptibility gene (15–19). Germline missense and nonsense point mutations were found in four of five CD families, one of these families also having LDD (18). These findings were confirmed by two other groups (20,21). Further analyses were also able to implicate *PTEN* as the susceptibility gene for BZS, with germline nonsense and missense point mutations being identified in two BZS families (22).

In this study, we have performed *PTEN* mutation analysis in 37 CD families and seven BZS families. Genotype–phenotype association analyses were carried out within the CD group in an attempt to identify correlations that may have clinical implications in the management of these patients in the future.

RESULTS

PTEN mutation analysis in CD and BZS

Constitutive DNA was collected from affected and non-affected members of CD and BZS families and analysed for *PTEN* mutations. Germline *PTEN* mutations were identified in 30 of 37 (81%) CD families (Table 1). These mutations were scattered over the entire gene with the exception of the first, fourth and last exons (Figs 1 and 2). Where additional family members, both affected and non-affected, were available, mutations were shown to segregate with the disease phenotype in CD families. None of the missense mutations identified was observed in a panel of 100 control alleles.

Table 1. *PTEN* mutations in CD and BZS

Family identifier	Mutation	Exon/IVS
CD families		
F97009 (LDD)	c.158insATAC	exon 2
F97019	c.158delTA	exon 2
CDmy	IVS2–2A→G	IVS2
F97017	I67R	exon 3
F92003 (LDD)	IVS3 +5G→A	IVS3
CDay	Q87X	exon 5
CDve	c.302–304del(TCA)insCC	exon 5
F96078	c.304insT	exon 5
CDcl	Q110X	exon 5
CDst	c.347–351del	exon 5
CDht; CDme	C124R	exon 5
CD2053; CDbh	G129E	exon 5
CDpl; CDtm	R130X	exon 5
CDuy	R130L	exon 5
CDwt (LDD)	E157X	exon 5
CDte	G165V (likely splice)	exon 6
F97013	c.565delA	exon 6
CDbn; CDsi	R233X	exon 7
F93036	c.723ins TT	exon 7
F96006	Q245X	exon 7
F95044	c.800insA	exon 7
CDfn	c.971insAT	exon 7
CDne	Y315X	exon 8
F95016	c.981delA	exon 8
CD93	R335X	exon 8
F93019	c.1007insA	exon 8
CD0014	no mutation	–
cdln	no mutation	–
CDma	no mutation	–
CDkk	no mutation	–
CDwr	no mutation	–
F91006	no mutation	–
F93034	no mutation	–
BZS families		
BanRO	Y68H	exon 3
BanR	S170R	exon 6
BanO	c.520–544del	exon 6
BanS	R233X	exon 7
BanSm-1	no mutation	–
BanSm-2	no mutation	–
BanW	no mutation	–

Of the 30 CD *PTEN* mutations identified, 10 (33%) were nonsense point mutations, six (20%) were missense point mutations, six (20%) were insertions, four (13%) were deletions, three (10%) were splice site mutations and one (3%) was a deletion/insertion mutation (Figs 1 and 2). Thirteen of 30 (43%) mutations were located in exon 5, which encodes the protein tyrosine phosphatase (PTPase) core motif at residues 122–132, IHCKAGKGRTG (Fig. 2). Exons 7 and 8, believed to encode potential tyrosine kinase phosphorylation sites at amino acids 233–240 (exon 7) and 308–315 (exon 8), contained six of 30 (20%) and four of 30 (13%) mutations, respectively. Together, exons 5, 7 and 8 contained 77% (23/30) of all the mutations identified in CD families.

Seven of the 13 mutations in exon 5 were within the core motif, five being missense point mutations. Of the 13 exon 5 mutations, five were 5′ of the core motif, and one occurred 3′ of the core motif. It is worthy of note that all 13 of the mutations falling 3′ of the PTPase core motif are predicted to truncate the PTEN protein, whilst only 11 of the 17 (65%) falling 5′ and within this motif are predicted to truncate. In contrast, five of the six missense mutations were found in exon 5, all within the interval encoding the PTPase core motif (Fig. 2).

Only four mutations, C124R, G129E, R130X and R233X, were each shown to occur in more than one CD family (Table 1). Two of these four, R130X and R233X, occurred at CpG dinucelotides. Haplotype analyses in these families using microsatellite markers in the 20 cM interval flanking *PTEN* excluded founder effects for common *PTEN* mutations in the respective families (data not shown).

Germline *PTEN* mutations were identified in four of seven (57%) BZS families. Where available, additional family members, both affected and non-affected, were screened and mutations were shown to segregate with the disease phenotype. None of these mutations was found in the PTPase core motif. Two BZS families had mutations in exon 6, the first, a 25 bp deletion, c.520–544del25, predicted to cause premature truncation of the protein, and the second, a missense point mutation, S170R. Other BZS mutations were found in exon 3 (Y68H) and exon 7 (R233X) (Table 1). The S170R and R233X mutations have been published previously (22). Interestingly, the R233X exon 7 mutation is present in both a single BZS family and two unrelated CD families.

Hemizygote analysis at IVS8 +32 T/G

Potential hemizygosity at the *PTEN* locus in mutation-negative families was assessed by screening for a T/G polymorphism within *PTEN* intron 8 detected by differential digestion with the restriction endonuclease *Hin*cII. This polymorphism is moderately heterozygous, with an earlier study finding 50% of samples to be informative for this marker (23). In three of the seven CD families without detected germline *PTEN* mutations, there was retention of heterozygosity at this marker in at least one affected member, thus excluding the possibility of whole gene deletion in these families. The three BZS families without detectable germline *PTEN* mutation were homozygous at this site.

Genotype–phenotype analyses in CD

Thirty five of the 37 CD families screened for *PTEN* mutations were analysed for genotype–phenotype correlations. Two CD families, CDne and F97017, did not have adequate clinical data to be included in genotype–phenotype studies. For two other previously published families, CD0014 (mutation-negative) and CD2053 (mutation-positive), clinicians could not come to an agreement on the nature (benign or malignant) of CNS involvement, and genotype–phenotype analyses were performed both with and without these families: these did not affect the results of analyses. Phenotypic data including CD2053 are presented (Tables 2, 3 and 4; Figs 1 and 2). Of the families with detected mutations, GI, skin and thyroid sites were most commonly involved, and also the most likely sites to have only benign tumours (Table 2).

Given the relatively small sample size, $n = 35$ for breast involvement and $n = 25$ for number of organ sites involved, we

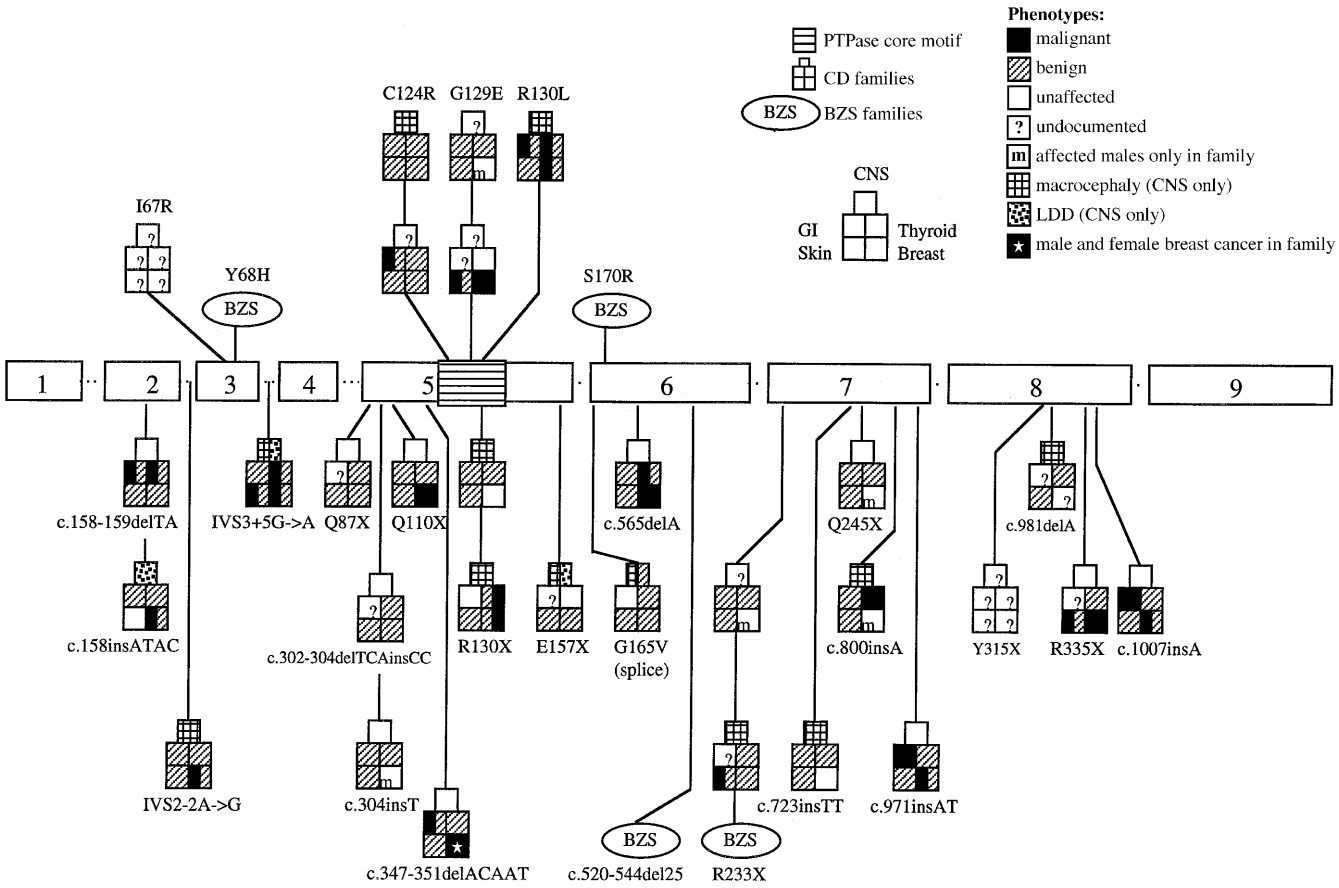

Figure 1. *PTEN* mutations in CD and BZS and their corresponding phenotypes. The PTPase core motif is marked in exon 5. CD families and BZS families are marked on figure. Phenotypes are as indicated on figure. Intervening sequence or splice site mutations are marked IVS, with the exception of G165V occurring in the donor splice site of exon 6. Truncating-type mutations are indicated in the lower half of the figure, and non-truncating (missense) mutations in the upper half.

had little power to detect even moderately large associations at a significance level of $P = 0.05$. Therefore, in order not to miss potential clinically relevant correlations and in accordance with standard statistical protocol given our small sample size, we chose to screen our data using $P = 0.10$ as a cut off to define possibly interesting genotype–phenotype associations for future examination in independent analyses. Of the many comparisons tested, we found associations between (i) the presence of a detectable mutation and breast involvement (unaffected, benign or malignant) (Table 3), and (ii) an interdependent association between the number of organ sites involved and the position and nature of the mutation (Table 4).

The breast phenotype–*PTEN* mutation-positive association might have potential clinical relevance (Table 3). Of the seven mutation-negative families, one had malignant breast disease and six had benign breast involvement. Among the 27 CD families who were mutation positive and have clinical breast information, seven did not have breast involvement, eight had benign breast disease and 12 had adenocarcinomas of the breast. This represents a significant difference between the categories (unaffected versus benign versus malignant) assessed ($P = 0.05$). Specifically and more directly, an association was also observed between the presence of a *PTEN* mutation and malignant breast disease ($P = 0.08$).

The possible association between the number of organ sites involved in a given family and the position and/or type of mutation may also be of clinical relevance (Table 4). Mutations 5′ of the PTPase core motif and within the motif appear to be associated with maximum organ involvement (5 versus ≤4 sites) ($P = 0.12$). The majority, five of seven, of the mutations within the core motif were missense (the only other missense mutation occurring in exon 3), so it is perhaps not surprising that we observed an interdependent association between non-truncating mutations and five-organ involvement ($P = 0.11$).

No obvious genotype–phenotype associations were noted for the three families with CD and LDD, nor were any correlations observed between thyroid disease and mutation type or location. It is of note that male breast cancer, previously not recognized as a component of CD, was observed in a single CD family, CDst.

DISCUSSION

Affected members from 30 of 37 (81%) CD families and four of seven (57%) BZS families were found to have germline *PTEN* mutations. *PTEN* truncating and missense mutations, associated with CD, were scattered along the entire gene, with the exception of the first, fourth and last exons.

In the seven of 37 (19%) CD families and three of seven (43%) BZS families where a *PTEN* mutation was not identified in the coding region or flanking intronic sequences, genetic heterogeneity or another mechanism of PTEN inactivation, could not be excluded.

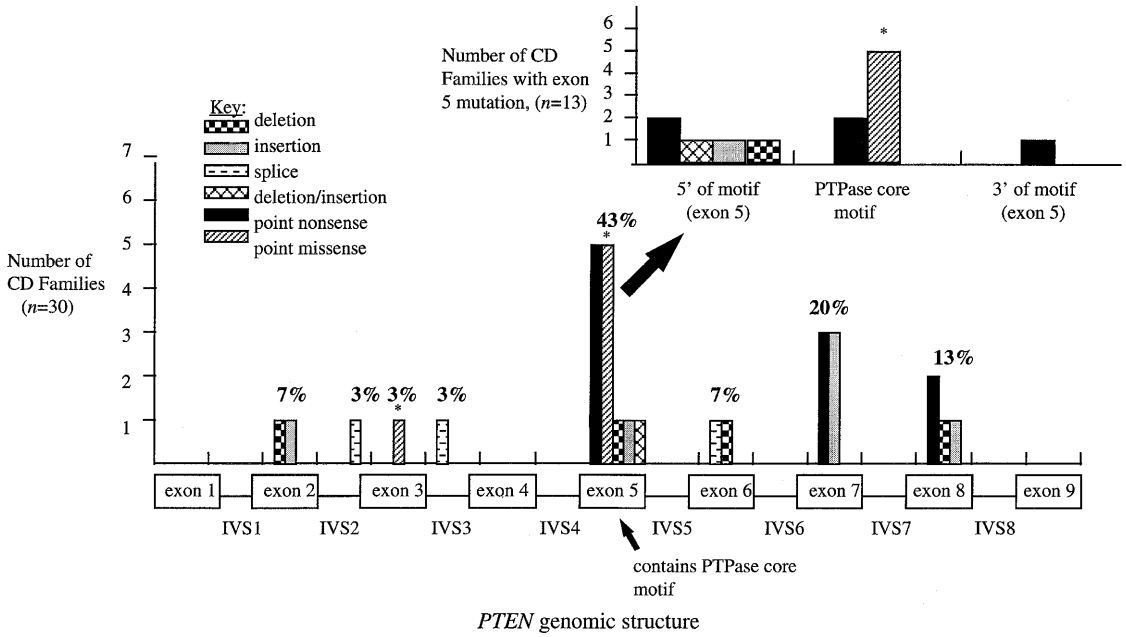

Figure 2. The distribution and nature of *PTEN* mutations in 30 CD families. Percentages reflect the frequency of a mutation in a given exon or intervening sequence (IVS) when compared with the gene as a whole (rounding error means that these percentages sum to 99%). Non-truncating mutations are denoted by an asterisk.

Hemizygote analyses in this study suggest that gross gene deletion is not a major mechanism in CD. If mutations were present within the promoter region or deep within an intron in these families, they would not have been detected by the methods used in this study. It is possible that transcriptional inactivation as the result of aberrant methylation at CpG islands in the as yet uncharacterized *PTEN* promoter region may result in monoallelic PTEN expression. However, to our knowledge, there is as yet no precedent for aberrant promoter methylation at CpG islands in the germline.

In this study, 30 distinct germline mutations were found in 35 CD and BZS families; four mutations, C124R, G129E, R130X and R233X, occurred twice each in CD, R233X also occurring in a patient with BZS. Two of these four, R130X and R233X, occurred at CpG dinucleotides. Haplotype analysis was able to exclude a common ancestor or founder effect in these latter families. Interestingly, R233X was also identified in a single BZS family as well as in two unrelated CD families. Again, haplotype sharing among these three families was excluded. Thus, an identical mutation would appear to be causing both the CD and BZS phenotype. Given this small series of mutation-positive BZS families and that only one BZS mutation, R233X, is common with CD, there are two broad possibilities. First, this BZS case is a misdiagnosis and is, in fact, CD. With the current clinical data on this family (22), this is probably excluded. Second, these data reaffirm the hypothesis that despite CD and BZS being two distinct clinical entities they might represent a variable spectrum of the same primary disorder (22). One could postulate, in the absence of further functional data, that germline R233X is capable of acting in a differential tissue-specific manner in these two syndromes. Alternatively, and perhaps more likely, one cannot exclude the possibility of other genetic modulatory factors or epigenetic factors being present. One such example of this is the putative modifying locus identified for familial adenomatous polyposis coli (FAP) mapped to chromosome 1p35–p36 (24). Identical germline mutations in *APC* can predispose an individual to FAP, characterized by adenomatous polyposis and colon cancer, Gardner syndrome, which has the additional features of mandibular osteomas and desmoids, and a proportion of Turcot syndrome, characterized by polyposis, colon cancer and medulloblastoma (25).

Mutations which cluster in certain exons could result from one or more factors, such as base pair composition [e.g. CpG dinucleotides (26)] or size of the exon, or they could reflect the biology of the protein product, i.e. important functional or structural domains. Among the CD families, 77% (23/30) of all mutations occurred in exons 5, 7 and 8. The great majority, 43%, occurred in exon 5. Compared with the rest of the gene, this is not particularly GC-rich and has only two CpG dinucleotides, one of which is mutated in three CD families—R130L in one family and R130X in two families. Although exon 5 is the largest exon, it represents 20% of the coding region of this gene, while 43% of CD mutations occurred in this exon. Hence, the clustering of almost half of all CD mutations here is likely to be of some functional significance. Exon 5 encodes the PTPase core motif, residues 122–132, IHCKAGKGRTG. It is worthy of note that none of the four *PTEN* mutations associated with BZS occurred in this motif. Interestingly, the majority of CD mutations occurring within the core motif itself (five of seven) were missense mutations, implying the sensitivity of this important functional domain to single amino acid changes.

After exon 5, a second 'hot spot' was observed in exons 7 and 8, believed to encode potential tyrosine kinase phosphorylation sites at amino acids 233–240 (exon 7) and 308–315 (exon 8). Exon 7 contained six of 30 (20%) and exon 8, four of 30 (13%). The frequency of mutations in these two exons (33% of all CD mutations) might also reflect their length (30% of the coding sequence).

Table 2. Phenotype characterization of CD families with germline *PTEN* mutations

Category A	
Any malignancy	15
No malignancy	6
Undocumented[a]	7
Total	28
Category B	
Four or fewer sites	21
All five sites	4
Undocumented[b]	3
Total	28
Category C	

Disease site/ phenotype	Unaffected	Benign only	Malignant	Macrocephaly only[c]	LDD/any tumour[c]	Undocumented	Total
CNS[c]	11	n/a	n/a	9	4	4	28
Thyroid	2	20	6	n/a	n/a	0	28
Skin	1	23	4	n/a	n/a	0	28
GI	2	13	6	n/a	n/a	7	28
Breast	7	8	12	n/a	n/a	1	28

[a]Undocumented in category A indicates cases where the phenotype is a combination of benign and undocumented organ sites, thus the possibility of an undocumented site being a malignancy cannot be excluded.
[b]Undocumented in category B indicates families where non-documentation at one or more sites means they cannot be classified as ≤4 or 5 organ sites affected.
[c]Phenotypes for CNS are 'unaffected', 'macrocephaly only' or 'LDD/any tumour'. For all other sites, phenotypes are 'benign only', 'malignant' or 'unaffected'. These classifications are not applicable (n/a) for other organ sites.

While it may be obvious that truncating mutations could disrupt normal protein function, the significance of missense mutations is difficult to assess in the absence of functional data. The possible functional role of a small number of CD and BZS *PTEN* missense mutations has been explored *in vitro* (19). For example, the missense point mutation S170R (exon 6), found in the BZS family BanR, was shown to dramatically decrease the phosphatase activity of PTEN (19). It might be expected that missense mutations in the PTPase motif act by disrupting PTPase function. Surprisingly, the missense point mutation G129E, identified in two unrelated CD families, was not found to decrease phosphatase activity, whereas mutation of the same residue to arginine, G129R, does decrease phosphatase activity (19). It is plausible that the G129E mutation identified in CD patients functions to alter substrate specificity, disrupt protein stability or disrupt RNA stability, in a manner that remains to be elucidated.

Another interesting missense mutation identified in this study was found in the cysteine in the PTPase motif, residue 124 of PTEN. This cysteine previously has been shown to be essential for catalysis, acting as a nucleophile that attacks the phosphorus atom in the phosphate moiety of its substrate, forming a thiol-phosphate intermediate (27). It has also been shown previously that mutation of this cysteine to alanine or serine will ablate phosphatase activity (28). Two CD families, CDme and CDht, were found to have missense mutation of this cysteine to arginine, C124R, which we postulate may also affect phosphatase activity.

Among CD families, two possible genotype–phenotype associations with some clinical relevance were found. The first was a significant correlation between *PTEN* mutation-positive CD families compared with mutation-negative families and the three categories of breast involvement (unaffected versus benign versus malignant breast disease). If PTEN in these mutation-negative families were to be inactivated by alternative mechanisms such as methylation or promoter mutation, then the described correlation would suggest that such inactivation would be associated with a milder breast phenotype. Alternatively, whilst genetic heterogeneity has yet to be observed in CD (14), the possibility that these families are linked to a locus other than *PTEN* which is less likely to cause breast malignancy cannot be excluded (21).

Table 3. Relationships between breast involvement and the presence or absence of *PTEN* mutations in CD families

Genotype	Breast phenotype			
	Unaffected	Affected	Benign	Malignant
Mutation	7	20	8	12
No mutation	0	7 (P = 0.30)	6	1 (P = 0.08)

P = 0.05 for a 3 × 2 table comparing *PTEN* mutation-negative versus *PTEN* mutation-positive with three different categories of breast involvement (unaffected versus benign versus malignant).

A second significant interdependent association was observed between non-truncating, i.e. missense, mutations and the involvement of all five organ sites (CNS, thyroid, breast, skin and GI tract) in CD. This association is perhaps not surprising when we consider that the majority of missense mutations, five of six, are located within the PTPase core motif that is likely to be of major functional significance. At present, we can only speculate as to the functional affect of these mutations. It may be possible that the mutant protein translated from an allele with a missense mutation interacts with a wild-type molecule, thus sequestering

wild-type protein and drastically reducing the level of normal PTEN protein available for phosphatase signalling. More plausibly, molecules with missense mutations may still interact with their substrates, but phosphatase activity is either reduced or absent. Hence, these mutant molecules sequester normal PTEN substrate away from wild-type PTEN. In contrast, truncated PTEN might be postulated to be unable to interact with its normal substrates and so does not sequester substrate from wild-type protein encoded by the wild-type allele. Given these postulates, we could surmise that sequestration of substrates, hence the presence of more phosphorylated substrates, could lead to multi-organ involvement. Alternatively, PTEN may function as a classic tumour suppressor following Knudson's 'two-hit' model of tumorigenesis (29), where the absence of both normal copies of the gene leads to tumour formation, or, in the case of CD and BZS, possibly hamartomas, in affected tissue. In this case, the association we have observed between the presence of missense mutations and maximum organ involvement may reflect a true positional effect of these missense mutations being clustered in the PTPase core motif of PTEN. Confirmation of these theories awaits further functional and molecular epidemiological studies of PTEN.

Table 4. Relationship between number of organ sites involved and specific *PTEN* mutation type or location in CD families

	Number of organ sites involved		
	≤4 sites	5 sites	
Mutation type			
Point missense	2	2	
Point nonsense	8	0	
Frameshift	10	0	
Splice site	1	2	$P = 0.004$
Total	21	4	
Truncating	19	2	
Non-truncating	2	2	$P = 0.11$
Total	21	4	
Splice site	2	2	
Non-splice site	19	2	$P = 0.11$
Total	21	4	
Mutation location			
5′ of and including the PTPase motif	11	4	
3′ of PTPase motif	10	0	$P = 0.12$
Total	21	4	

No additional CD genotype–phenotype associations were observed in this study, possibly due to the modest size of this series. A few interesting observations are notable. The three CD families reported with LDD were shown to have germline *PTEN* mutation at three separate sites: two splice site mutations in IVS2 and IVS3 and a truncating mutation in exon 5, downstream of the PTPase core motif. It is of note that the four patients co-segregating CD and LDD reported by Nelen *et al.* (20) were also found to have only splice site or truncating mutations. One other interesting phenotype observed is that of male breast cancer that previously has not been recognized as a component of CD. One CD family with a germline deletion in exon 5, c.347–351del, generating a stop codon at c.372, amino acid 124, within the PTPase core motif, has a single case of male invasive ductal carcinoma of the breast. Recently, one other case of male breast involvement has been reported in a CD family with a germline *PTEN* mutation; however, it is unclear whether this involvement is of a benign or malignant nature (20). Further studies should be performed to confirm these suggestions that male breast cancer is part of the CD spectrum.

The identification of germline mutations in *PTEN* in CD and BZS will allow DNA-based predictive testing, the results of which can be incorporated into the clinical management of these families. This data set, the largest presented to date, has already suggested a number of interesting associations in CD families between the presence of breast malignancy and a germline *PTEN* mutation, and a possibly interdependent association of the presence of germline missense mutations clustered in the core motif and 5′ of it and the involvement of all organ sites affected in CD. However, confirmation of these associations and the identification of new genotype–phenotype correlations, especially in BZS, awaits the study of a larger data set.

MATERIALS AND METHODS

Patients

Affected and non-affected members from 37 CD families were collected for *PTEN* mutation analysis. The diagnoses of CD were made in accordance with the criteria of the International Cowden Consortium (4,14). Three CD families, F97009, F92003 and CDwt, also had LDD. Detailed clinical information was unavailable for two families—CDne and F97017. Since original submission of the clinical phenotype for CD0014 and CD2053, clinicians for these families have resubmitted opposing clinical features (e.g. malignant versus benign). These discrepancies in clinical features do not detract from the fact that these families are classic CD. However, genotype–phenotype analyses had to be performed twice, one with one set of features and one with the other.

Seven BZS families were also collected for *PTEN* mutation analysis. Whilst international consensus criteria have not yet been established for the diagnosis of BZS given the small number of families reported in the literature, samples were ascertained on the basis of the presence of macrocephaly, lipomatosis and pigmented macules on the glans penis.

DNA preparation

Blood samples were obtained from affected and non-affected members of 37 CD and seven BZS families. Constitutional DNA was extracted from blood leucocytes using standard techniques (30).

PTEN mutation analysis

The nine exons of *PTEN*, including flanking intronic sequence, were sequenced using primers designed from intronic sequences. Primers and PCR conditions have been previously described (15,18) with the exception of primers for exons 2 and 4, respectively P10X2AF102 (5′-GTT TGA TTG CTG CAT ATT TCA G-3′) and P10X2AR88 (5′-TCT AAA TGA AAA CAC AAC ATG-3′), P10X2CF51 (5′-CAT TAT AAA GAT TCA GGC AAT G-3′) and P10X2CR108 (5′-GAC AGT AAG ATA CAG

TCT ATC-3′). Purification of PCR products was carried out by gel isolation followed by purification using the Wizard PCR Preps DNA Purification System (Promega, Madison, WI). Purified PCR products were sequenced directly using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Norwalk, CT). Cycle sequencing products were electrophoresed on 6% Long Ranger gels (FMC Bioproducts, Rockland, ME) and analysed on an Applied Biosystems model 373A automated DNA sequencer (Perkin-Elmer).

Haplotype generation and hemizygote analysis

Haplotype analyses were performed to exclude a possible founder effect in CD families with identical germline *PTEN* mutations. The seven microsatellite markers used were from the 20 cM interval flanking *PTEN*, *D10S219* (AFM240wg7)–*D10S551* (AFM 240vf10)–*D10S579* (AFM282yc1)–*D10S541* (AFM205xe3)–*D10S1739* (AFMb362yg5)–*D10S564* (AFM029xh12)–*D10S583* (AFM289zh5) (markers ordered centromeric to telomeric, reflecting the integrated genetic and physical map of chromosome 10q, http://www-genome.wi.mit.edu). PCR conditions for these markers have been described previously (31,32).

In an attempt to exclude gross deletion of *PTEN* in CD and BZS families without germline *PTEN* mutation, a polymorphism, IVS8 +32G/T, was screened for by differential digestion with *Hin*cII under the manufacturer's conditions (New England Biolabs, Beverly, MA).

Genotype–phenotype analysis methods

Genotype–phenotype analyses were performed on a family-as-a-unit basis (33,34). Potential intrafamilial correlations (individual-as-unit analysis) were unable to be assessed given that in most families samples from multiple affected members have yet to be collected. Phenotype data were not available for two families–F97107 and CDne. Mutations were categorized in several alternative ways. First, families with a detected mutation were compared with those with no detected mutation. Among families with mutations, we considered mutation type; point missense, point nonsense, frameshifts (combining deletions, insertions and the deletion/insertion mutation) and splice site, truncating versus non-truncating and splice site versus non-splice site. Mutation location defined by 5′ of and including the PTPase core motif or 3′ of the PTPase motif was also analysed.

Each genotype coding was cross-tabulated with the phenotype of five potential sites of organ involvement, specifically the CNS, skin, thyroid, breast and GI, using the family as the unit of analysis. Two alternative codings for each of the five sites of involvement were considered. For the CNS, we used a three-category coding (not affected, macrocephaly only, and tumour and/or LDD) and a two-category coding (not affected, affected). For skin, thyroid, breast and GI phenotypes, we used a three-category coding (not affected, benign only, malignant) and a two-category coding (not affected, affected). We also examined two composite phenotypes: malignancy at any organ site versus no malignancy, and number of organ sites involved (operationally defined as all five organ sites versus four or fewer sites). Three families, CDbn, CDht and F95016, were not included in this last analysis as non-documentation at one or more sites could not exclude the possibility of all sites being affected.

For all tables, the statistical significance of differences between genotype groups were computed using Fisher's two-tailed exact test (35,36).

ACKNOWLEDGEMENTS

We would like to thank the families for participating in this study. We are grateful to A. Satterfied and S. Hew for assistance with sample collection and data management; K. Schneider, S. Keiffer and A. Chittenden for seeing CD families with C.E.; and S. Cummings, M. Gadd, A.M. Goldstein, J.J. Mulvihill, M.A. Tucker and A. Rustgi for assistance with CD individuals. Dr S. Verselis is thanked for helpful discussions. The Molecular Biology Core Facility at the Dana-Farber Cancer Institute, Boston, is acknowledged for running sequencing and Genescan gels. This study was supported by the American Cancer Society (RPG 97-064-01VM) (to C.E.); the Susan G. Komen Foundation (to P.L.M.D. and C.E.); Women's Cancer Program (Dana-Farber Cancer Institute); Barr Investigatorship (to C.E.); l'Institut Bergonié, La Ligue Nationale Française Contre le Cancer and Les Comités Départementaux de la Corréze et de la Gironde de la Ligue Contre le Cancer (to M.L.). C.E. is the Lawrence and Susan Marx Investigator in Human Cancer Genetics.

REFERENCES

1. Carlson, H.E., Burns, T.W., Davenport, S.L., Luger, A.M., Spence, M.A., Sparkes, R.S. and Orth, D.N. (1984) Cowden disease: gene marker studies and measurements of epidermal growth factor. *Am. J. Hum. Genet.*, **38**, 908–917.
2. Hanssen, A.M.N. and Fryns, J.P. (1995) Cowden syndrome. *J. Med. Genet.*, **32**, 117–119.
3. Longy, M. and Lacombe, D. (1996) Cowden disease. Report of a family and review. *Ann. Genet.*, **39**, 35–42.
4. Eng, C. (1997) Cowden syndrome. *J. Genet. Counsel.*, **6**, 181–191.
5. Starink, T.M., van der Veen, J.P.W., Arwert, F., de Waal, L.P., de Lange, G.G., Gille, J.J.P. and Eriksson, A.W. (1986) The Cowden syndrome: a clinical and genetic study in 21 patients. *Clin. Genet.*, **29**, 222–233.
6. Padberg, G.W., Schot, J.D., Vielvoye, G.J., Bots, G.T. and de Beer, F.C. (1991) Lhermitte–Duclos disease and Cowden disease: a single phakomatosis. *Ann. Neurol.*, **29**, 517–523.
7. Albrecht, S., Haber, R.M., Goodman, J.C. and Duvic, M. (1992) Cowden syndrome and Lhermitte–Duclos disease. *Cancer*, **70**, 869–876.
8. Vinchon, M., Blond, S., Lejeune, J.P., Krivosik, I., Fossati, P., Assaker, R. and Christiaens, J.L. (1994) Association of Lhermitte–Duclos and Cowden disease: report of a new case and review of the literature. *J. Neurol., Neurosurg. Psychiatr.*, **57**, 699–704.
9. Eng, C., Murday, V., Seal, S., Mohammed, S., Hodgson, S.V., Chaudary, M.A., Fentiman, I.S., Ponder, B.A. and Eeles, R.A. (1994) Cowden syndrome and Lhermitte–Duclos disease in a family: a single genetic syndrome with pleiotropy? *J. Med. Genet.*, **31**, 458–461.
10. Bannayan, G.A. (1971) Lipomatosis, angiomatosis, and macrencephalia. A previously undescribed congenital syndrome. *Arch. Pathol.*, **92**, 1–5.
11. Gorlin, R.J., Cohen, M.M., Condon, L.M. and Burke, B.A. (1992) Bannayan–Riley–Ruvalcaba syndrome. *Am. J. Med. Genet.*, **44**, 307–314.
12. Zonana, J., Rimoin, D.L. and Davis, D.C. (1976) Macrocephaly with multiple lipomas and hemangiomas. *J. Paediatr.*, **89**, 600–603.
13. Fargnoli, M.C., Orlow, S.J., Semelconcepcion, J. and Bolognia, J.L. (1996) Clinicopathologic findings in the Bannayan–Riley–Ruvalcaba syndrome. *Arch. Dermatol.*, **132**, 1214–1218.
14. Nelen, M.R., Padberg, G.W., Peeters, E.A.J., Lin, A.Y., van den Helm, B., Frants, R.R., Coulon, V., Goldstein, A.M., van Reen, M.M.M., Easton, D.F., Eeles, R.A., Hodgson, S., Mulvihill, J.J., Murday, V.A., Tucker, M.A., Mariman, E.C.M., Starink, T.M., Ponder, B.A.J., Ropers, H.H., Kremer, H., Longy, M. and Eng, C. (1996) Localization of the gene for Cowden disease to chromosome 10q22–23. *Nature Genet.*, **13**, 114–116.

15. Steck, P.A., Pershouse, M.A., Jasser, S.A., Slfred Yung, W.K., Lin, H., Ligon, A.H., Langford, L.A., Baumgard, M.L., Hattier, T., Davis, T., Frye, C., Hu, R., Swedlund, B., Teng, D.H.F. and Tavtigian, S.V. (1997) Identification of a candidate tumour suppressor gene, *MMAC1*, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nature Genet*., **15**, 356–362.
16. Li, D.-M. and Sun, H. (1997) TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor β. *Cancer Res*., **57**, 2124–2129.
17. Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S., Puc, J., Miliaresis, C., Rodgers, L., McCombie, R., Bigner, S.H., Giovanella, B.C., Ittman, M., Tyoko, B., Hibshoosh, H., Wigler, M.H. and Parsons, R. (1997) *PTEN*, a putative protein tyrosine phosphatase gene mutated in human brain, breast and prostate cancer. *Science*, **275**, 1943–1947.
18. Liaw, D., Marsh, D.J., Li, J., Dahia, P.L.M., Wang, S.I., Zheng, Z., Bose, S., Call, K.M., Tsou, H.C., Peacocke, M., Eng, C. and Parsons, R. (1997) Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature Genet*., **16**, 64–67.
19. Myers, M.P., Stolarov, J.P., Eng, C., Li, J., Wang, S.I., Wigler, M.H., Parsons, R. and Tonks, N.K. (1997) P-TEN, the tumor suppressor from human chromosome 10q23, is a dual specificity phosphatase. *Proc. Natl Acad. Sci. USA*, **94**, 9052–9057.
20. Nelen, M.R., van Staveren, W.C.G., Peeters, E.A.J., Hassel, M.B., Gorlin, R.J., Hamm, H., Lindboe, C.F., Fryns, J.-P., Sijmons, R.H., Woods, D.G., Mariman, E.C.M., Padberg, G.W. and Kremer, H. (1997) Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. *Hum. Mol. Genet*., **6**, 1383–1387.
21. Tsou, H.C., Teng, D., Ping, X.I., Broncolini, V., Davis, T., Hu, R., Xie, X.-X., Gruener, A.C., Schrager, C.A., Christiano, A.M., Eng, C., Steck, P., Ott, J., Tavitigian, S.V. and Peacocke, M. (1997) The role of *MMAC1* mutations in early-onset breast cancer: causative in association with Cowden syndrome and excluded in *BRCA1*-negative cases. *Am. J. Hum. Genet.*, **61**, 1036–1043.
22. Marsh, D.J., Dahia, P.L.M., Zheng, Z., Parsons, R., Gorlin, R.J. and Eng, C. (1997) Germline mutations in *PTEN* are present in Bannayan–Zonana syndrome. *Nature Genet*., **16**, 333–334.
23. Dahia, P.L.M., Marsh, D.J., Zheng, Z., Zedenius, J., Komminoth, P., Frisk, T., Wallin, G., Parsons, R., Longy, M., Larsson, C. and Eng, C. (1997) Somatic deletions and mutations in the Cowden disease gene, *PTEN*, in sporadic thyroid tumors. *Cancer Res*., **57**, 4710–4713.
24. Tomlinson, I.P.M., Neale, K., Talbot, I.C., Spigelman, A.D., Williams, C.B., Phillips, R.K.S. and Bodmer, W.F. (1996) A modifying locus for familial adenomatous polyposis may be present on chromosome 1p35–36. *J. Med. Genet*., **33**, 268–273.
25. Rustgi, A.K. (1994) Medical progress—hereditary gastrointestinal polyposis and nonpolyposis syndromes. *N. Engl. J. Med*., **331**, 1694–1702.
26. Lindahl, T. (1974) An *N*-glycosidase from *Escherichia col*i that relaeases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl Acad. Sci. USA*, **71**, 3649–3653.
27. Barford, D., Flint, A.J. and Tonks, N.K. (1994) Crystal structure of human protein tyrosine phosphatase 1B. *Science*, **263**, 1397–1404.
28. Flint, A.J., Tiganis, T., Barford, D. and Tonks, N.K. (1997) Development of 'substrate-trapping' mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc. Natl Acad. Sci. USA*, 94, 1680–1685.
29. Knudson, A.G. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA*, **68**, 820–823.
30. Mathew, C.G., Smith, B.A., Thorpe, K., Wong, Z., Royle, N.J., Jeffreys, A.J. and Ponder, B.A.J. (1987) Deletion of genes on chromosome 1 in endocrine neoplasia. *Nature*, **328**, 524–526.
31. Marsh, D.J., Zheng, Z., Zedenius, J., Kremer, H., Padberg, G.W., Larsson, C., Longy, M. and Eng, C. (1997) Differential loss of heterozygosity in the region of the Cowden locus within 10q22–23 in follicular thyroid adenomas and carcinomas. *Cancer Res*., **57**, 500–503.
32. Marsh, D.J., Dahia, P.L.M., Coulon, V., Zheng, Z., Dorion-Bonnet, F., Call, K.M., Little, R., Lin, A.Y., Eeles, R.A., Goldstein, A.M., Hodgson, S.V., Richardson, A.-L., Robinson, B.G., Weber, H.C., Longy, M. and Eng, C. (1998) Allelic imbalance, including deletion of *PTEN/MMAC1*, at the Cowden disease locus on 10q22–23, in hamartomas from patients with Cowden syndrome and germline *PTEN* mutation. *Genes Chrom. Cancer*, in press.
33. Mulligan, L.M., Marsh, D.J., Robinson, B.G., Schuffenecker, I., Zedenius, J., Lips, C.J.M., Gagel, R.F., Takai, S.I., Noll, W.W., Niederle, B., Raue, F., Lacroix, A., Thibodeau, S.N., Frilling, A., Ponder, B.A.J. and Eng, C. for the International *RET* Mutation Consortium (1995) Genotype–phenotype correlation in MEN 2: report of the International *RET* Mutation Consortium. *J. Int. Med*., **238**, 343–246.
34. Eng, C., Clayton, D., Schuffenecker, I., Lenoir, G., Cote, G., Gagel, R.F., Ploos van Amstel, H.K., Lips, C.J.M., Nishisho, I., Takai, S.-I., Marsh, D.J., Robinson, B.G., Frank-Raue, K., Raue, F., Xue, F., Noll, W.W., Romei, C., Pacini, F., Fink, M., Niederle, B., Zedenius, J., Nordenskjöld, M., Komminoth, P., Hendy, G.N., Gharib, H., Thibodeau, S.N., Lacroix, A., Frilling, A., Ponder, B.A.J. and Mulligan, L.M. (1996) The relationship between specific RET proto-oncogene mutations and disease phenotype in multiple endocrine neoplasia type 2. *J. Am. Med. Assoc*., **276**, 1575–1579.
35. Mehta, C.R. and Patel, N.R. (1983) A network algorithm for performing Fisher's exact test in rXc contingency tables. *J Am. Statistical Assoc*., **78**, 427–434.
36. Agresti, A. (1990) *Categorical Data Analysis*. John Wiley and Sons, New York.